

Efficient approximation of the solution of certain nonlinear reaction–diffusion equation I: the case of small absorption[☆]

Ezequiel Dratman

*Instituto de Ciencias, Universidad Nacional de General Sarmiento, Juan M. Gutiérrez
1150 (B1613GSX) Los Polvorines, Buenos Aires, Argentina.*

Abstract

We study the positive stationary solutions of a standard finite-difference discretization of the semilinear heat equation with nonlinear Neumann boundary conditions. We prove that, if *the absorption is small enough*, compared with the flux in the boundary, there exists a unique solution of such a discretization, which approximates the unique positive stationary solution of the “continuous” equation. Furthermore, we exhibit an algorithm computing an ε -approximation of such a solution by means of a homotopy continuation method. The cost of our algorithm is *linear* in the number of nodes involved in the discretization and the logarithm of the number of digits of approximation required.

Keywords: Two-point boundary-value problem, finite differences, Neumann boundary condition, stationary solution, homotopy continuation, polynomial system solving, condition number, complexity
2000 MSC: 65H10, 65L10, 65L12, 65H20, 65Y20

[☆]The research was partially supported by the following grants: UNGS 30/3084, UNGS 30/1066, CIC (2007–2009) and PIP 11220090100421.

Email address: edratman@ungs.edu.ar (Ezequiel Dratman)

URL: <http://sites.google.com/site/ezequieldratman> (Ezequiel Dratman)

1. Introduction

This article deals with the following semilinear heat equation with Neumann boundary conditions:

$$\begin{cases} u_t = u_{xx} - g_1(u) & \text{in } (0, 1) \times [0, T), \\ u_x(1, t) = \alpha g_2(u(1, t)) & \text{in } [0, T), \\ u_x(0, t) = 0 & \text{in } [0, T), \\ u(x, 0) = u_0(x) \geq 0 & \text{in } [0, 1], \end{cases} \quad (1)$$

where $g_1, g_2 \in \mathcal{C}^3(\mathbb{R})$ are analytic functions in $x = 0$ and α is a positive constant. The nonlinear heat equation models many physical, biological and engineering phenomena, such as heat conduction (see, e.g., [1, §20.3], [2, §1.1]), chemical reactions and combustion (see, e.g., [3, §5.5], [4, §1.7]), growth and migration of populations (see, e.g., [5, Chapter 13], [2, §1.1]), etc. In particular, “power-law” nonlinearities have long been of interest as a tractable prototype of general polynomial nonlinearities (see, e.g., [3, §5.5], [6, Chapter 7], [7], [8], [2, §1.1]).

The long-time behavior of the solutions of (1) has been intensively studied (see, e.g., [9], [10], [11], [12], [13], [14], [15], [16] and the references therein). In order to describe the dynamic behavior of the solutions of (1) it is usually necessary to analyze the behavior of the corresponding *stationary solutions* (see, e.g., [13], [9]), i.e., the positive solutions of the following two-point boundary-value problem:

$$\begin{cases} u_{xx} = g_1(u) & \text{in } (0, 1), \\ u_x(1) = \alpha g_2(u(1)), \\ u_x(0) = 0. \end{cases} \quad (2)$$

The usual numerical approach to the solution of (1) consists of considering a second-order finite-difference discretization in the variable x , with a uniform mesh, keeping the variable t continuous (see, e.g., [17]). This semi-discretization in space leads to the following initial-value problem:

$$\begin{cases} u'_1 = \frac{2}{h^2}(u_2 - u_1) - g_1(u_1), \\ u'_k = \frac{1}{h^2}(u_{k+1} - 2u_k + u_{k-1}) - g_1(u_k), & (2 \leq k \leq n-1) \\ u'_n = \frac{2}{h^2}(u_{n-1} - u_n) - g_1(u_n) + \frac{2\alpha}{h}g_2(u_n), \\ u_k(0) = u_0(x_k), & (1 \leq k \leq n) \end{cases} \quad (3)$$

where $h := 1/(n-1)$ and x_1, \dots, x_n define a uniform partition of the interval $[0, 1]$. A similar analysis to that in [18] shows the convergence of the positive

solutions of (3) to those of (1) and proves that every bounded solution of (3) tends to a stationary solution of (3), namely to a solution of

$$\begin{cases} 0 &= \frac{2}{h^2}(u_2 - u_1) - g_1(u_1), \\ 0 &= \frac{1}{h^2}(u_{k+1} - 2u_k + u_{k-1}) - g_1(u_k), \quad (2 \leq k \leq n-1) \\ 0 &= \frac{2}{h^2}(u_{n-1} - u_n) - g_1(u_n) + \frac{2\alpha}{h}g_2(u_n). \end{cases} \quad (4)$$

Hence, the dynamic behavior of the positive solutions of (3) is rather determined by the set of solutions $(u_1, \dots, u_n) \in (\mathbb{R}_{>0})^n$ of (4).

Very little is known concerning the study of the stationary solutions of (3) and the comparison between the stationary solutions of (3) and (1). In [13], [18] and [19] there is a complete study of the positive solutions of (4) for the particular case $g_1(x) := x^p$ and $g_2(x) := x^q$, i.e., a complete study of the positive solutions of

$$\begin{cases} 0 &= \frac{2}{h^2}(u_2 - u_1) - u_1^p, \\ 0 &= \frac{1}{h^2}(u_{k+1} - 2u_k + u_{k-1}) - u_k^p, \quad (2 \leq k \leq n-1) \\ 0 &= \frac{2}{h^2}(u_{n-1} - u_n) - u_n^p + \frac{2\alpha}{h}u_n^q. \end{cases} \quad (5)$$

In [13] it is shown that there are spurious solutions of (4) for $q < p < 2q - 1$, that is, positive solutions of (4) not converging to any solution of (2) as the mesh size h tends to zero.

In [18] and [19] there is a complete study of (4) for $p > 2q - 1$ and $p < q$. In these articles it is shown that in such cases there exists exactly one positive real solution. Furthermore, a numeric algorithm solving a given instance of the problem under consideration with $n^{O(1)}$ operations is proposed. In particular, the algorithm of [19] has linear cost in n , that is, this algorithm gives a numerical approximation of the desired solution with $O(n)$ operations.

We observe that the family of systems (5) has typically an exponential number $O(p^n)$ of *complex* solutions ([20]), and hence it is ill conditioned from the point of view of its solution by the so-called robust universal algorithms (cf. [21], [22], [23]). An example of such algorithms is that of general continuation methods (see, e.g., [24]). This shows the need of algorithms specifically designed to compute positive solutions of “structured” systems like (4).

Continuation methods aimed at approximating the real solutions of non-linear systems arising from a discretization of two-point boundary-value problems for second-order ordinary differential equations have been con-

sidered in the literature (see, e.g., [25], [26], [27]). These works are usually concerned with Dirichlet problems involving an equation of the form $u_{xx} = f(x, u, u_x)$ for which the existence and uniqueness of solutions is known. Further, they focus on the existence of a suitable homotopy path rather on the cost of the underlying continuation algorithm. As a consequence, they do not seem to be suitable for the solution of (4). On the other hand, it is worth mentioning the analysis of [28] on the complexity of shooting methods for two-point boundary value problems.

Let $g_1, g_2 \in \mathcal{C}^3(\mathbb{R})$ be analytic functions in $x = 0$ such that $g_i(0) = 0$, $g'_i(x) > 0$, $g''_i(x) > 0$ and $g'''_i(x) \geq 0$ for all $x > 0$ with $i = 1, 2$. We observe that g_1 and g_2 are a wide generalization of the monomial functions of system (5). Moreover, we shall assume throughout the paper that the function $g := g_1/g_2$ is strictly decreasing, generalizing thus the relation $p < q$ in (5). In this article we study the existence and uniqueness of the positive solutions of (4), and we obtain numerical approximations of these solutions using homotopy methods. In a forthcoming paper, we shall consider the generalization of the relations $q < p < 2q - 1$ and $2q - 1 < p$ in (5).

1.1. Our contributions

In the first part of the article we prove that (4) has a unique positive solution, and we obtain upper and lower bounds for this solution independent of h , generalizing the results of [19].

In the second part of the article we exhibit an algorithm which computes an ε -approximation of the positive solution of (4). Such an algorithm is a continuation method that tracks the positive real path determined by the smooth homotopy obtained by considering (4) as a family of systems parametrized by α . Its cost is roughly of $n \log \log \varepsilon$ arithmetic operations, improving thus the exponential cost of general continuation methods.

The cost estimate of our algorithm is based on an analysis of the condition number of the corresponding homotopy path, which might be of independent interest. We prove that such a condition number can be bounded by a quantity independent of $h := 1/n$. This in particular implies that each member of the family of systems under consideration is significantly better conditioned than both an “average” dense system (see, e.g., [29, Chapter 13, Theorem 1]) and an “average” sparse system ([30, Theorem 1]).

1.2. Outline of the paper

Section 2 is devoted to determine the number of positive solutions of (4). For this purpose, we prove that the homotopy of systems mentioned

above is smooth (Theorem 5). From this result we deduce the existence and uniqueness of the positive solutions of (4).

In Section 3 we obtain upper and lower bounds for the coordinates of the positive solution of (4).

In Section 4 we obtain estimates on the condition number of the homotopy path considered in the previous section (Theorem 14). Such estimates are applied in Section 5 in order to estimate the cost of the homotopy continuation method for computing the positive solution of (4).

2. Existence and uniqueness of stationary solutions

Let U_1, \dots, U_n be indeterminates over \mathbb{R} . Let g_1 and g_2 be two functions of class $\mathcal{C}^3(\mathbb{R})$ such that $g_i(0) = 0$, $g'_i(x) > 0$, $g''_i(x) > 0$ and $g'''_i(x) \geq 0$ for all $x > 0$ with $i = 1, 2$. As stated in the introduction, we are interested in the positive solutions of (4) for a given positive value of α , that is, in the positive solutions of the nonlinear system

$$\begin{cases} 0 &= -(U_2 - U_1) + \frac{h^2}{2}g_1(U_1), \\ 0 &= -(U_{k+1} - 2U_k + U_{k-1}) + h^2g_1(U_k), \quad (2 \leq k \leq n-1) \\ 0 &= -(U_{n-1} - U_n)\frac{h^2}{2}g_1(U_n) - h\alpha g_2(U_n), \end{cases} \quad (6)$$

for a given value $\alpha = \alpha^* > 0$, where $h := 1/(n-1)$. Observe that, as α runs through all possible values in $\mathbb{R}_{>0}$, one may consider (6) as a family of nonlinear systems parametrized by α , namely,

$$\begin{cases} 0 &= -(U_2 - U_1) + \frac{h^2}{2}g_1(U_1), \\ 0 &= -(U_{k+1} - 2U_k + U_{k-1}) + h^2g_1(U_k), \quad (2 \leq k \leq n-1) \\ 0 &= -(U_{n-1} - U_n)\frac{h^2}{2}g_1(U_n) - hAg_2(U_n), \end{cases} \quad (7)$$

where A is a new indeterminate.

2.1. Preliminary analysis

Let A, U_1, \dots, U_n be indeterminates over \mathbb{R} , set $U := (U_1, \dots, U_n)$ and denote by $F : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ the nonlinear map defined by the right-hand side of (7). From the first $n-1$ equations of (7) we easily see that, for a given positive value $U_1 = u_1$, the (positive) values of U_2, \dots, U_n, A are uniquely

determined. Therefore, letting U_1 vary, we may consider U_2, \dots, U_n , A as functions of U_1 , which are indeed recursively defined as follows:

$$\begin{aligned} U_1(u_1) &:= u_1, \\ U_2(u_1) &:= u_1 + \frac{h^2}{2}g_1(u_1), \\ U_{k+1}(u_1) &:= 2U_k(u_1) - U_{k-1}(u_1) + h^2g_1(U_k(u_1)), \quad (2 \leq k \leq n-1), \\ A(u_1) &:= \left(\frac{1}{h}(U_n - U_{n-1})(u_1) + \frac{h}{2}g_1(U_n(u_1)) \right) / g_2(U_n(u_1)). \end{aligned} \tag{8}$$

Arguing recursively, one deduces the following lemma (cf. [20, Remark 20]).

Lemma 1. *For any $u_1 > 0$, the following assertions hold:*

- i. $(U_k - U_{k-1})(u_1) = h^2 \left(\frac{1}{2}g_1(u_1) + \sum_{j=2}^{k-1} g_1(U_j(u_1)) \right) > 0$,
- ii. $U_k(u_1) = u_1 + h^2 \left(\frac{k-1}{2}g_1(u_1) + \sum_{j=2}^{k-1} (k-j)g_1(U_j(u_1)) \right) > 0$,
- iii. $(U'_k - U'_{k-1})(u_1) = h^2 \left(\frac{1}{2}g'_1(u_1) + \sum_{j=2}^{k-1} g'_1(U_j(u_1))U'_j(u_1) \right) > 0$,
- iv. $U'_k(u_1) = 1 + h^2 \left(\frac{k-1}{2}g'_1(u_1) + \sum_{j=2}^{k-1} (k-j)g'_1(U_j(u_1))U'_j(u_1) \right) > 1$,

for $2 \leq k \leq n$.

For the proof of the next lemma we need the following technical result

Remark 2. *Let $f_1, f_2, f_3 \in \mathcal{C}^2(\mathbb{R}_{>0})$ be positive functions such that*

- $f''_1(x) > 0$,
- $f'_2(x) > 0$, $f'_3(x) > 0$,
- $f_2(x) > f_3(x)$,

for all $x > 0$. Let $F : \mathbb{R}_{>0} \rightarrow \mathbb{R}$ be the function defined by

$$F(x) := \frac{f_1(f_2(x)) - f_1(f_3(x))}{f_2(x) - f_3(x)}.$$

Then $F'(x) > 0$ for all $x > 0$.

Proof. Fix $x > 0$. By the definition of F we have that

$$\begin{aligned} F'(x)(f_2(x) - f_3(x)) &= \left(f'_1(f_2(x))f'_2(x) - f'_1(f_3(x))f'_3(x) \right) \\ &\quad - F(x)(f'_2(x) - f'_3(x)) \end{aligned}$$

holds. From the Mean Value Theorem, there exists $\xi \in (f_3(x), f_2(x))$ with $F(x) = f'_1(\xi)$. Therefore

$$F'(x)(f_2(x) - f_3(x)) = \left(f'_1(f_2(x)) - f'_1(\xi)\right)f'_2(x) + \left(f'_1(\xi) - f'_1(f_3(x))\right)f'_3(x).$$

Since f'_1 , f_2 and f_3 are strictly increasing functions, we conclude that $F'(x) > 0$. \square

Now we prove an important result for the existence and uniqueness of the solutions of (7)

Lemma 3. *For any $u_1 > 0$, the following assertions hold:*

- i. $\left(\frac{U_k - U_{k-1}}{g_1(U_k)}\right)'(u_1) < 0$,
- ii. $\left(\frac{U_k - U_1}{g_1(U_k)}\right)'(u_1) < 0$,
- iii. $\left(\frac{U_k - U_{k-1}}{U_k - U_1}\right)'(u_1) \geq 0$,
- iv. $\left(\frac{g_1(U_k)}{g_1(U_1)}\right)'(u_1) > 0$,

for $2 \leq k \leq n$.

Proof. Let $L_{j,i} : \mathbb{R}_{>0} \rightarrow \mathbb{R}$ be the function defined by

$$L_{j,i}(u_1) := \frac{g_1(U_j) - g_1(U_i)}{U_j - U_i}(u_1),$$

where $1 \leq i < j \leq n$. From Remark 2 and Lemma 1, we deduce that

$$L'_{j,i}(u_1) = \left(\frac{g_1(U_j) - g_1(U_i)}{U_j - U_i}\right)'(u_1) > 0. \quad (9)$$

By (8) we have

$$\begin{aligned} \frac{U_2 - U_1}{g_1(U_2)}(u_1) &= \left(\frac{2}{h^2} + L_{2,1}(u_1)\right)^{-1}, \\ \frac{g_1(U_2)}{g_1(U_1)}(u_1) &= 1 + \frac{h^2}{2}L_{2,1}(u_1). \end{aligned}$$

Combining these identities with (9) we obtain (i), (ii), (iii) and (iv) for $k = 2$. Now, arguing inductively, suppose that our statement is true for a given $k \geq 2$. From (8) we have:

$$\begin{aligned} \left(\frac{U_{k+1} - U_k}{U_{k+1} - U_1}\right)(u_1) &= \left(1 + \frac{U_k - U_1}{U_{k+1} - U_k}\right)^{-1}(u_1) \\ &= \left(1 + \left(\frac{U_k - U_{k-1}}{U_k - U_1} + \frac{g_1(U_k)h^2}{U_k - U_1}\right)^{-1}\right)^{-1}(u_1). \end{aligned}$$

Applying the inductive hypotheses we deduce that (iii) holds for $k + 1$. On the other hand, by Lemma 1 and (8) we have that

$$\begin{aligned}
\frac{U_{k+1}-U_k}{g_1(U_{k+1})}(u_1) &= \left(\frac{g_1(U_k)}{U_{k+1}-U_k} + L_{k+1,k} \right)^{-1}(u_1) \\
&= \left(\left(\frac{U_k-U_{k-1}}{g_1(U_k)} + h^2 \right)^{-1} + L_{k+1,k} \right)^{-1}(u_1), \\
\frac{U_{k+1}-U_1}{g_1(U_{k+1})}(u_1) &= \left(\frac{g_1(U_k)}{U_{k+1}-U_1} + L_{k+1,k} \frac{U_{k+1}-U_k}{U_{k+1}-U_1} \right)^{-1}(u_1) \\
&= \left(\left(\frac{U_k-U_{k-1}}{g_1(U_k)} + \frac{U_k-U_1}{g_1(U_k)} + h^2 \right)^{-1} + L_{k+1,k} \frac{U_{k+1}-U_k}{U_{k+1}-U_1} \right)^{-1}(u_1), \\
\frac{g_1(U_{k+1})}{g_1(U_1)}(u_1) &= \left(L_{k+1,1} \frac{U_{k+1}-U_1}{g_1(U_1)} \right)(u_1) + 1 \\
&= h^2 \left(L_{k+1,1} \left(\frac{k-1}{2} + \sum_{j=2}^k (k-j) \frac{g_1(U_j)}{g_1(U_1)} \right) \right)(u_1) + 1.
\end{aligned}$$

hold. Combining the inductive hypothesis with (9) and (iii), for $k + 1$, we conclude that (i), (ii) and (iv) hold for $k + 1$. \square

2.2. Existence and uniqueness

Let $P : (\mathbb{R}_{>0})^2 \rightarrow \mathbb{R}$ be the nonlinear map defined by

$$P(\alpha, u_1) := \frac{1}{h}(U_{n-1}(u_1) - U_n(u_1)) - \frac{h}{2}g_1(U_n(u_1)) + \alpha g_2(U_n(u_1)). \quad (10)$$

Observe that $P(\alpha, U_1) = 0$ represents the minimal equation satisfied by the coordinates (α, u_1) of any (complex) solution of the nonlinear system (7). Therefore, for fixed $\alpha > 0$, the positive roots of $P(\alpha, U_1)$ are the values of u_1 we want to obtain. Furthermore, from the parametrizations (8) of the coordinates u_2, \dots, u_n of a given solution $(\alpha, u_1, \dots, u_n) \in (\mathbb{R}_{>0})^{n+1}$ of (7) in terms of u_1 , we conclude that the number of positive roots of $P(\alpha, U_1)$ determines the number of positive solutions of (7) for such a value of α .

Therefore, we analyze the existence of positive roots of the function $P(\alpha, U_1)$ for values $\alpha > 0$. Let $g : \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}$ be the function defined by

$$g(x) := \frac{g_1}{g_2}(x). \quad (11)$$

By Lemma 1(i) we have that

$$\begin{aligned}
P(\alpha, u_1) &= \alpha g_2(U_n(u_1)) - h \left(\frac{1}{2}g_1(u_1) + \sum_{j=2}^{n-1} g_1(U_j(u_1)) + \frac{1}{2}g_1(U_n(u_1)) \right) \\
&\geq \alpha g_2(U_n(u_1)) - g_1(U_n(u_1)) = g_2(U_n(u_1)) \left(\alpha - g(U_n(u_1)) \right)
\end{aligned}$$

holds for any $u_1 > 0$.

Suppose that g is surjective. Therefore, there exist $u_1^*, u_1^{**} > 0$ such that $g(U_n(u_1^*)) = \alpha$ and $g(U_n(u_1^{**})) = 2\alpha/h$ hold. From this choice of u_1^* and u_1^{**} and the inequality above we deduce

$$\begin{aligned} P(\alpha, u_1^*) &\geq g_2(U_n(u_1^*))(\alpha - g(U_n(u_1^*))) = 0, \\ P(\alpha, u_1^{**}) &= \frac{1}{h}(U_{n-1}(u_1^{**}) - U_n(u_1^{**})) + g_2(U_n(u_1^{**}))\left(\alpha - \frac{h}{2}g(U_n(u_1^{**}))\right) \\ &= \frac{1}{h}(U_{n-1}(u_1^{**}) - U_n(u_1^{**})) \leq 0. \end{aligned}$$

Since $P(A, U_1)$ is a continuous function in $(\mathbb{R}_{>0})^2$, from the previous considerations we obtain the following result.

Proposition 4. *Fix $\alpha > 0$ and $n \in \mathbb{N}$. If the function g of (11) is surjective, then (7) has a positive solution with $A = \alpha$.*

In order to establish the uniqueness, we prove that the homotopy path that we obtain by moving the parameter α in $\mathbb{R}_{>0}$ is smooth. For this purpose, we show that the rational function $A(U_1)$ implicitly defined by the equation $P(A, U_1) = 0$ is decreasing. We observe that an explicit expression for this function in terms of U_1 is obtained in (8).

Theorem 5. *Let $A(U_1)$ be the rational function of (8). If the function g of (11) is decreasing, then the condition $A'(u_1) < 0$ is satisfied for every $u_1 \in \mathbb{R}_{>0}$.*

Proof. Let U_1, U_2, \dots, U_n, A be the functions defined in (8). Observe that A can be rewritten as follows:

$$A = g(U_n) \left(\frac{U_n - U_{n-1}}{hg_1(U_n)} + \frac{h}{2} \right).$$

Taking derivatives with respect to U_1 , we have

$$A' = g'(U_n)U_n' \left(\frac{U_n - U_{n-1}}{hg_1(U_n)} + \frac{h}{2} \right) + g(U_n) \left(\frac{U_n - U_{n-1}}{hg_1(U_n)} \right)'.$$

Fix $u_1 > 0$. From Lemma 1 we see that $g(U_n)(u_1)$ is positive. Furthermore, by Lemma 3, we have

$$\left(\frac{U_n - U_{n-1}}{hg_1(U_n)} \right)'(u_1) < 0.$$

These remarks show that

$$A'(u_1) < \left(g'(U_n) U'_n \left(\frac{U_n - U_{n-1}}{h g_1(U_n)} + \frac{h}{2} \right) \right)(u_1) = \left(g'(U_n) U'_n \frac{A}{g(U_n)} \right)(u_1). \quad (12)$$

From Lemma 1, we deduce that $U'_n(u_1)$ and $A(u_1)$ are positive. Combining this affirmation with the monotonicity of g we deduce the statement of the theorem. \square

Now we state and prove the main result of this section.

Theorem 6. *Let be given $\alpha > 0$ and $n \in \mathbb{N}$. If the function g of (11) is surjective and decreasing, then (6) has a unique positive solution.*

Proof. Proposition 4 shows that (6) has solutions in $(\mathbb{R}_{>0})^n$ for any $\alpha > 0$ and any $n \in \mathbb{N}$. Therefore, there remains to show the uniqueness assertion.

By Theorem 5, the condition $A'(u_1) < 0$ holds for every $u_1 \in \mathbb{R}_{>0}$. Arguing by contradiction, assume that there exist two distinct positive solutions $(u_1, \dots, u_n), (\hat{u}_1, \dots, \hat{u}_n) \in (\mathbb{R}_{>0})^n$ of (6) for α . This implies that $u_1 \neq \hat{u}_1$ and $A(u_1) = A(\hat{u}_1)$, where $A(U_1)$ is defined in (8). But this contradicts the fact that $A'(u_1) < 0$ holds in $\mathbb{R}_{>0}$, showing thus the theorem. \square

3. Bounds for the positive solution

In this section we obtain bounds for the positive solution of (7). More precisely, we find an interval containing the positive solution of (7) whose endpoints only depend on α . These bounds will allow us to establish an efficient procedure of approximation of this solution.

Lemma 7. *Let $(\alpha, u) \in (\mathbb{R}_{>0})^{n+1}$ be a solution of (7) for $A = \alpha$. Then*

$$\alpha g_2(u_n) < g_1(u_n).$$

Proof. From the last equation of (7) for $A = \alpha$ and Lemma 1(i), we obtain the identity

$$\alpha g_2(u_n) = h \left(\frac{1}{2} g_1(u_1) + g_1(u_2) + \dots + g_1(u_{n-1}) + \frac{1}{2} g_1(u_n) \right).$$

From the previous identity and Lemma 1(i) we immediately deduce the statement of the lemma. \square

From Lemma 7 we obtain the following corollary.

Corollary 8. *Let $(\alpha, u) \in (\mathbb{R}_{>0})^{n+1}$ be a solution of (7) for $A = \alpha$. If the function g of (11) is surjective and strictly decreasing, then*

$$u_n < g^{-1}(\alpha).$$

Let $(\alpha, u) \in (\mathbb{R}_{>0})^{n+1}$ be a solution of (7) for $A = \alpha$. In the following lemma we obtain an upper bound of u_n in terms of u_1 and α .

Lemma 9. *Let $(\alpha, u) \in (\mathbb{R}_{>0})^{n+1}$ be a solution of (7) for $A = \alpha$. If the function g of (11) is surjective and strictly decreasing, then $u_n < e^M u_1$ holds, with $M := g'_1(g^{-1}(\alpha))$.*

Proof. Let $(\alpha, u) \in (\mathbb{R}_{>0})^{n+1}$ be a solution of (7) for $A = \alpha$. Combining Lemma 1(i) and the Mean Value Theorem, we obtain the following identities

$$\begin{aligned} u_{k+1} &= u_k + h^2 \left(\frac{g_1(u_1)}{2} + g_1(u_2) + \cdots + g_1(u_k) \right) \\ &= u_k + h^2 \left(\frac{g'_1(\xi_1)u_1}{2} + g'_1(\xi_2)u_2 + \cdots + g'_1(\xi_k)u_k \right) \end{aligned}$$

for $1 \leq k \leq (n-1)$, where $\xi_i \in [0, u_i]$ for $1 \leq i \leq k$. Since g'_1 is an increasing function in $\mathbb{R}_{>0}$, combining Lemma 1(i) and Corollary 8, we obtain

$$\begin{aligned} u_{k+1} &\leq u_k + h^2(g'_1(u_1)u_1 + \cdots + g'_1(u_k)u_k) \\ &\leq (1 + hg'_1(g^{-1}(\alpha)))u_k = (1 + hM)u_k \end{aligned}$$

for $1 \leq k \leq (n-1)$. Arguing recursively, we deduce that

$$u_n \leq (1 + hM)^{n-1} u_1 \leq e^M u_1.$$

This completes the proof. \square

In our next lemma we obtain a lower bound of u_1 in terms of α .

Lemma 10. *Let $(\alpha, u) \in (\mathbb{R}_{>0})^{n+1}$ be a solution of (7) for $A = \alpha$. If the function g of (11) is surjective and strictly decreasing, then*

$$u_1 > g^{-1}(\alpha C(\alpha))$$

holds, where $C(\alpha) \geq 1$ is a constant such that

$$\lim_{\alpha \rightarrow +\infty} C(\alpha) = 1.$$

Proof. From Lemma 9 and Lemma 1(i) we deduce the inequalities

- $g_2(u_n) < g_2(e^M u_1)$,
- $g_1(u_1) < h\left(\frac{1}{2}g_1(u_1) + g_1(u_2) + \cdots + g_1(u_{n-1}) + \frac{1}{2}g_1(u_n)\right) = \alpha g_2(u_n)$.

with $M := g'_1(g^{-1}(\alpha))$. Combining both inequalities we obtain

$$g_1(u_1) < \alpha g_2(e^M u_1) = \alpha \frac{g_2(e^M u_1)}{g_2(u_1)} g_2(u_1). \quad (13)$$

Since g_2 is an analytic function in $x = 0$ and $g_2(x) \neq 0$ for every $x > 0$, exists $k \geq 1$ such that

$$\lim_{x \rightarrow 0^+} \frac{g_2(e^M x)}{g_2(x)} = e^{kM}.$$

Combining this with Corollary 8, we deduce that there exists $C(\alpha) > 0$ such that

$$1 \leq \frac{g_2(e^M u_1)}{g_2(u_1)} \leq C(\alpha).$$

Furthermore, we can choose $C(\alpha)$ with

$$\lim_{\alpha \rightarrow +\infty} C(\alpha) = 1.$$

Combining this remark with (13) we obtain

$$g_1(u_1) < \alpha C(\alpha) g_2(u_1),$$

which immediately implies the statement of the lemma. \square

4. Numerical conditioning

Let be given $n \in \mathbb{N}$ and $\alpha^* > 0$. In order to compute the positive solution of (7) for this value of n and $A = \alpha^*$, we shall consider (7) as a family of systems parametrized by the values α of A , following the positive real path determined by (7) when A runs through a suitable interval whose endpoints are α_* and α^* , where α_* is a positive constant independent of h to be fixed in Section 5.

A critical measure for the complexity of this procedure is the condition number of the path considered, which is essentially determined by the inverse of the Jacobian matrix of (7) with respect to the variables U_1, \dots, U_n , and the gradient vector of (7) with respect to the variable A on the path. In this section we prove the invertibility of such a Jacobian matrix, and obtain an explicit form of its inverse. Then we obtain an upper bound on the condition number of the path under consideration.

4.1. The Jacobian matrix

Let $F := F(A, U) : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ be the nonlinear map defined by the right-hand side of (7). In this section we analyze the invertibility of the Jacobian matrix of F with respect to the variables U , namely,

$$J(A, U) := \frac{\partial F}{\partial U}(A, U) := \begin{pmatrix} \Gamma_1 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & \Gamma_n \end{pmatrix},$$

with $\Gamma_1 := 1 + \frac{1}{2}h^2g'_1(U_1)$, $\Gamma_i := 2 + h^2g'_1(U_i)$ for $2 \leq i \leq n-1$ and $\Gamma_n := 1 + \frac{1}{2}h^2g'_1(U_n) - hAg'_2(U_n)$.

We start relating the nonsingularity of the Jacobian matrix $J(\alpha, u)$ with that of the corresponding point in the path determined by (7). Let $(\alpha, u) \in (\mathbb{R}_{>0})^{n+1}$ be a solution of (7) for $A = \alpha$. Taking derivatives with respect to U_1 in (8) and substituting u_1 for U_1 we obtain the following tridiagonal system:

$$\begin{pmatrix} \Gamma_1(u_1) & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & \Gamma_n(u_1) \end{pmatrix} \begin{pmatrix} 1 \\ U'_2(u_1) \\ \vdots \\ U'_n(u_1) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ hg_2(U_n(u_1))A'(u_1) \end{pmatrix}.$$

For $1 \leq k \leq n-1$, we denote by $\Delta_k := \Delta_k(A, U)$ the k th principal minor of the matrix $J(A, U)$, that is, the $(k \times k)$ -matrix formed by the first k rows and the first k columns of $J(A, U)$. By the Cramer rule we deduce the identities:

$$hg_2(U_n(u_1))A'(u_1) = \det(J(\alpha, u)), \quad (14)$$

$$\det(J(\alpha, u))U'_k(u_1) = hg_2(U_n(u_1))A'(u_1) \det(\Delta_{k-1}(\alpha, u)), \quad (15)$$

for $2 \leq k \leq n$. Suppose that $\alpha > 0$ and that the function g of (11) is decreasing. Then Theorem 5 asserts that $A'(u_1) < 0$ holds. Combining this inequality with (14) we conclude that $\det(J(\alpha, u)) < 0$ holds. Furthermore, by (15), we have

$$U'_k(u_1) = \det(\Delta_{k-1}(\alpha, u)) \quad (2 \leq k \leq n). \quad (16)$$

Combining Remark 1(iv) and (16) it follows that $\det(\Delta_k(\alpha, u)) > 0$ holds for $1 \leq k \leq n-1$. As a consequence, we have that all the principal minors of the symmetric matrix $\Delta_{n-1}(\alpha, u)$ are positive. Then the Sylvester criterion shows that $\Delta_{n-1}(\alpha, u)$ is positive definite. These remarks allow us to prove the following result.

Theorem 11. *Let $(\alpha, v) \in (\mathbb{R}_{>0})^{n+1}$ be a solution of (7) for $A = \alpha$. If the function g of (11) is decreasing, then the matrix $J(\alpha, u)$ is invertible with $\det(J(\alpha, u)) < 0$. Furthermore, their $(n-1)$ th principal minor is symmetric and positive definite.*

Having shown the invertibility of the matrix $J(\alpha, u)$ for every solution $(\alpha, u) \in (\mathbb{R}_{>0})^{n+1}$ of (7), the next step is to obtain an explicit expression for the corresponding inverse matrices $J^{-1}(\alpha, u)$. For this purpose, we establish a result on the structure of the matrix $J^{-1}(\alpha, u)$.

Proposition 12. *Let $(\alpha, u) \in (\mathbb{R}_{>0})^{n+1}$ be a solution of (7). If the function g of (11) is decreasing, then the following matrix factorization holds:*

$$J^{-1}(\alpha, u) = \begin{pmatrix} 1 & \frac{1}{u_2} & \frac{1}{u_3} & \cdots & \frac{1}{u_n} \\ & 1 & \frac{u'_2}{u_3} & \cdots & \frac{u'_2}{u_n} \\ & & \ddots & \ddots & \vdots \\ & & & 1 & \frac{u'_{n-1}}{u_n} \\ & & & & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{u_2} & & & & \\ \frac{1}{u_3} & \frac{u'_2}{u_3} & & & \\ \vdots & \vdots & \ddots & & \\ \frac{1}{u_n} & \frac{u'_2}{u_n} & \cdots & \frac{u'_{n-1}}{u_n} & \\ \frac{1}{d(J)} & \frac{u'_2}{d(J)} & \cdots & \frac{u'_{n-1}}{d(J)} & \frac{u'_n}{d(J)} \end{pmatrix},$$

where $d(J) := \det(J(\alpha, u))$ and $u'_k := U'_k(u_1)$ for $2 \leq k \leq n$.

Proof. Since $J(\alpha, u)$ is symmetric, invertible, tridiagonal and their $(n-1)$ th principal minor is positive definite, the proof follows by a similar argument to that of [18, Proposition 25]. \square

4.2. Upper bounds on the condition number

From the explicit expression of the inverse of the Jacobian matrix $J(A, U)$ on the points of the real path determined by (7), we can finally obtain estimates on the condition number of such a path.

Let $\alpha^* > 0$ and $\alpha_* > 0$ be given constants independent of h . Suppose that the function g of (11) is surjective and decreasing. Then Theorem 6 proves that (7) has a unique positive solution with $A = \alpha$ for every α in the real interval $\mathcal{I} := \mathcal{I}(\alpha_*, \alpha^*)$ whose endpoints are α_* and α^* , which we denote by $(u_1(\alpha), U_2(u_1(\alpha)), \dots, U_n(u_1(\alpha)))$. We bound the condition number

$$\kappa := \max\{\|\varphi'(\alpha)\|_\infty : \alpha \in \mathcal{I}\},$$

associated to the function $\varphi : \mathcal{I} \rightarrow \mathbb{R}^n$, $\varphi(\alpha) := (u_1(\alpha), U_2(u_1(\alpha)), \dots, U_n(u_1(\alpha)))$.

For this purpose, from the Implicit Function Theorem we have

$$\begin{aligned}\|\varphi'(\alpha)\|_\infty &= \left\| \left(\frac{\partial F}{\partial U}(\alpha, \varphi(\alpha)) \right)^{-1} \frac{\partial F}{\partial A}(\alpha, \varphi(\alpha)) \right\|_\infty \\ &= \left\| J^{-1}(\alpha, \varphi(\alpha)) \frac{\partial F}{\partial A}(\alpha, \varphi(\alpha)) \right\|_\infty.\end{aligned}$$

We observe that $(\partial F / \partial A)(\alpha, \varphi(\alpha)) = (0, \dots, 0, -hg_2(U_n(u_1(\alpha))))^t$ holds. From Proposition 12 we obtain

$$\|\varphi'(\alpha)\|_\infty = \left\| \frac{hg_2(U_n(u_1(\alpha)))}{\det(J(\alpha, \varphi(\alpha)))} (1, U_2'(u_1(\alpha)), \dots, U_n'(u_1(\alpha)))^t \right\|_\infty.$$

Combining this identity with (14), we conclude that

$$\|\varphi'(\alpha)\|_\infty = \left\| \frac{1}{A'(u_1(\alpha))} (1, U_2'(u_1(\alpha)), \dots, U_n'(u_1(\alpha)))^t \right\|_\infty.$$

From Lemma 1, we deduce the following proposition.

Proposition 13. *Let $\alpha^* > 0$ and $\alpha_* > 0$ be given constants independent of h . Suppose that the function g of (11) is surjective and decreasing. Then*

$$\|\varphi'(\alpha)\|_\infty = \frac{U_n'(u_1(\alpha))}{|A'(u_1(\alpha))|}$$

holds for $\alpha \in \mathcal{I}$.

Combining Proposition 13 and (12) we conclude that

$$\|\varphi'(\alpha)\|_\infty < \frac{g(U_n(u_1(\alpha)))}{\alpha |g'(U_n(u_1(\alpha)))|}.$$

Applying Lemma 10 and Corollary 8 we deduce the following result.

Theorem 14. *Let $\alpha^* > 0$ and $\alpha_* > 0$ be given constants independent of h . Suppose that the function g of (11) is surjective and $g'(x) < 0$ holds for all $x \in \mathbb{R}_{>0}$. Then there exists a constant $\kappa_1(\alpha_*, \alpha^*) > 0$ independent of h with*

$$\kappa < \kappa_1(\alpha_*, \alpha^*).$$

5. An efficient numerical algorithm

As a consequence of the well conditioning of the positive solutions of (7), we shall exhibit an algorithm computing the positive solution of (7) for $A = \alpha^*$. This algorithm is a homotopy continuation method (see, e.g., [31, §10.4], [29, §14.3]) having a cost which is *linear* in n .

There are two different approaches to estimate the cost of our procedure: using Kantorovich-type estimates as in [31, §10.4], and using Smale-type estimates as in [29, §14.3]. We shall use the former, since we are able to control the condition number in suitable neighborhoods of the real paths determined by (7). Furthermore, the latter does not provide significantly better estimates.

Let $\alpha_* > 0$ be a constant independent of h . Suppose that the following conditions hold:

- g is surjective,
- $g'(x) < 0$ holds for all $x > 0$,
- $g''(x) \geq 0$ holds for all $x > 0$,

where g is the function of (11). Then the path defined by the positive solutions of (7) with $\alpha \in [\alpha^*, \alpha_*]$ is smooth, and the estimate of Theorem 14 hold. Assume that we are given a suitable approximation $u^{(0)}$ of the positive solution $\varphi(\alpha_*)$ of (7) for $A = \alpha_*$. In this section we exhibit an algorithm which, on input $u^{(0)}$, computes an approximation of $\varphi(\alpha^*)$. We recall that φ denotes the function which maps each $\alpha > 0$ to the positive solution of (7) for $A = \alpha$.

From Corollary 8 and Lemma 10, we have that the coordinates of the positive solution of (7) tend to zero when α tends to infinity. Therefore, for α large enough, we obtain a suitable approximation of the positive solution (7) for $A = \alpha_*$, and we track the positive real path determined by (7) until $A = \alpha^*$.

In order to deal with a bounded set, we consider the change of variables $B := 1/A$. Then system (7) for $A = \alpha^*$ can be rewritten in terms of B as follows:

$$\begin{cases} 0 &= -(U_2 - U_1) + \frac{h^2}{2}g_1(U_1), \\ 0 &= -(U_{k+1} - 2U_k + U_{k-1}) + h^2g_1(U_k), & (2 \leq k \leq n-1), \\ 0 &= -(U_{n-1} - U_n) - hB^{-1}g_2(U_n) + \frac{h^2}{2}g_1(U_n) \end{cases} \quad (17)$$

for $B = \beta^*$, where $\beta^* := 1/\alpha^*$.

Let $0 < \beta_* < \beta^*$ be a constant independent of h to be determined. Fix $\beta \in [\beta_*, \beta^*]$. By Corollary 8 it follows that $\varphi(\alpha)$ is an interior point of the compact set

$$K_\beta := \{u \in \mathbb{R}^n : \|u\|_\infty \leq 2g^{-1}(1/\beta)\},$$

where $\varphi : [\beta_*, \beta^*] \rightarrow \mathbb{R}^n$ is the function which maps each $\beta \in [\beta_*, \beta^*]$ to the positive solution of (17) for $B = \beta$, namely

$$\varphi(\beta) := (u_1(\beta), \dots, u_n(\beta)) := (u_1(\beta), U_2(u_1(\beta)), \dots, U_n(u_1(\beta))).$$

First we prove that the Jacobian matrix $J_\beta(u) := (\partial F / \partial U)(\beta, u)$ is invertible in a suitable subset of K_β . Let $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^n$ be points with

$$\|u - \varphi(\beta)\|_\infty < \delta_\beta, \quad \|v - \varphi(\beta)\|_\infty < \delta_\beta,$$

where $\delta_\beta > 0$ is a constant to be determined. Note that if $\delta_\beta \leq g^{-1}(1/\beta)$ then $u \in K_\beta$ and $v \in K_\beta$. By the Mean Value Theorem, we see that the entries of the diagonal matrix $J_\beta(u) - J_\beta(v)$ satisfy the estimates

$$\begin{aligned} |(J_\beta(u) - J_\beta(v))_{ii}| &\leq 2h^2 g_1''(2g^{-1}(1/\beta)) \delta_\beta \quad (1 \leq i \leq n-1), \\ |(J_\beta(u) - J_\beta(v))_{nn}| &\leq 2h \max\{g_2''(2g^{-1}(1/\beta))/\beta, g_1''(2g^{-1}(1/\beta))\} \delta_\beta. \end{aligned}$$

By Theorem 11 and Proposition 12 we have that the matrix $J_{\varphi(\beta)} := J_\beta(\varphi(\beta)) = (\partial F / \partial U)(\beta, \varphi(\beta))$ is invertible and

$$(J_{\varphi(\beta)}^{-1})_{ij} = \sum_{k=\max\{i,j\}}^{n-1} \frac{U'_i(u_1(\beta))U'_j(u_1(\beta))}{U'_k(u_1(\beta))U'_{k+1}(u_1(\beta))} + \frac{U'_i(u_1(\beta))U'_j(u_1(\beta))}{U'_n(u_1(\beta))\det(J_{\varphi(\beta)})}$$

holds for $1 \leq i, j \leq n$. According to Lemma 1, we have $U'_n(u_1(\beta)) \geq \dots \geq U'_2(u_1(\beta)) \geq 1$. These remarks show that

$$\begin{aligned} \|J_{\varphi(\beta)}^{-1}(J_\beta(u) - J_\beta(v))\|_\infty &\leq \\ &\leq \eta_\beta \delta_\beta \left(2 + \frac{h^2 + \sum_{j=2}^{n-1} h^2 U'_j(u_1(\beta)) + h U'_n(u_1(\beta))}{|\det(J_{\varphi(\beta)})|} \right) \quad (18) \\ &\leq 2\eta_\beta \delta_\beta \left(1 + \frac{h U'_n(u_1(\beta))}{|\det(J_{\varphi(\beta)})|} \right), \end{aligned}$$

where $\eta_\beta := 2 \max\{g_1''(2g^{-1}(1/\beta)), g_2''(2g^{-1}(1/\beta))/\beta\}$. Since $B = 1/A$, from Theorem 5 we deduce that $B'(u_1) = -A'(u_1)/A^2(u_1) > 0$ for all $x > 0$. Combining this assertion with (14), we obtain the following identity:

$$\frac{hU'_n(u_1(\beta))}{|\det(J_{\varphi(\beta)})|} = \frac{U'_n(u_1(\beta))B^2(u_1(\beta))}{B'(u_1(\beta))g_2(u_n(\beta))}.$$

From (12), we have that

$$\frac{hU'_n(u_1(\beta))}{|\det(J_{\varphi(\beta)})|} = \frac{U'_n(u_1(\beta))B^2(u_1(\beta))}{B'(u_1(\beta))g_2(u_n(\beta))} \leq \frac{g(u_n(\beta))B(u_1(\beta))}{|g'(u_n(\beta))|g_2(u_n(\beta))}. \quad (19)$$

Combining this inequality with the definition and the monotonicity of g , we deduce

$$\begin{aligned} \frac{|g'(u_n(\beta))|g_2(u_n(\beta))}{g(u_n(\beta))} &= \frac{g_1(u_n(\beta))g_2'(u_n(\beta)) - g_1'(u_n(\beta))g_2(u_n(\beta))}{g_1(u_n(\beta))} \\ &= g_2'(u_n(\beta)) \left(1 - \frac{g_1'(u_n(\beta))g_2(u_n(\beta))}{g_1(u_n(\beta))g_2'(u_n(\beta))} \right). \end{aligned}$$

Since g_1 and g_2 are analytic functions in $x = 0$ and $g_1(0) = g_2(0) = 0$, there exists $r > 0$ such that, for all $|x| < r$, we have that

$$g_1(x) = \sum_{k=p}^{\infty} c_k x^k, \quad g_2(x) = \sum_{k=q}^{\infty} d_k x^k,$$

with $c_p \neq 0$ and $d_q \neq 0$, where p and q are positive integers greater than 1. Hence, the following identity holds:

$$\lim_{x \rightarrow 0} \left(1 - \frac{g_1'(x)g_2(x)}{g_1(x)g_2'(x)} \right) = 1 - \frac{p}{q}.$$

Taking into account that $u_n(\beta) \in (0, g^{-1}(1/\beta^*))$ holds for all $\beta \in [\beta_*, \beta^*]$, we conclude

$$\frac{|g'(u_n(\beta))|g_2(u_n(\beta))}{g(u_n(\beta))} \geq g_2'(u_n(\beta))(1 - \rho^*),$$

where ρ^* is a constant which depends only on β^* . Combining the last inequality with (18) and (19), we obtain

$$\begin{aligned} \left\| J_{\varphi(\beta)}^{-1} \left(J_\beta(u) - J_\beta(v) \right) \right\|_\infty &\leq 2\eta_\beta \delta_\beta \left(1 + \frac{g(u_n(\beta))B(u_1(\beta))}{|g'(u_n(\beta))|g_2(u_n(\beta))} \right) \\ &\leq 2\eta_\beta \delta_\beta \left(1 + \frac{B(u_1(\beta))}{g_2'(u_n(\beta))(1 - \rho^*)} \right). \end{aligned}$$

From Lemma 10 and Corollary 8, there exists a constant $C(\beta) \geq 1$ such that

$$\left\| J_{\varphi(\beta)}^{-1} \left(J_{\beta}(u) - J_{\beta}(v) \right) \right\|_{\infty} \leq 2\eta_{\beta} \left(1 + \frac{\beta}{g_2'(g^{-1}(C(\beta)/\beta))(1-\rho^*)} \right) \delta_{\beta}. \quad (20)$$

Furthermore, the following condition holds:

$$\lim_{\beta \rightarrow 0^+} C(\beta) = 1.$$

Finally, since $g'(x) < 0$ holds for all $x > 0$, we have

$$\begin{aligned} \frac{\beta}{g_2'(g^{-1}(C(\beta)/\beta))(1-\rho^*)} &= \frac{C(\beta)}{g_2'(g^{-1}(C(\beta)/\beta))g(g^{-1}(C(\beta)/\beta))(1-\rho^*)} \\ &\leq \frac{C(\beta)}{g_1'(g^{-1}(C(\beta)/\beta))(1-\rho^*)}. \end{aligned}$$

Combining this inequality with (20), we deduce that

$$\left\| J_{\varphi(\beta)}^{-1} \left(J_{\beta}(u) - J_{\beta}(v) \right) \right\|_{\infty} \leq \frac{2\eta_{\beta}(\theta^* + 1)C(\beta)}{g_1'(g^{-1}(C(\beta)/\beta))(1-\rho^*)} \delta_{\beta}, \quad (21)$$

with $\theta^* := (1 - \rho^*)g_1'(g^{-1}(1/\beta^*))$. Hence, defining δ_{β} in the following way:

$$\delta_{\beta} := \min \left\{ \frac{g_1'(g^{-1}(C(\beta)/\beta))(1-\rho^*)}{8\eta_{\beta}(\theta^* + 1)C(\beta)}, g^{-1}(1/\beta) \right\}, \quad (22)$$

we obtain

$$\left\| J_{\varphi(\beta)}^{-1} \left(J_{\beta}(u) - J_{\beta}(v) \right) \right\|_{\infty} \leq \frac{1}{4}. \quad (23)$$

In particular, for $v = \varphi(\beta)$, this bound allows us to consider $J_{\beta}(u)$ as a perturbation of $J_{\varphi(\beta)}$. More precisely, by a standard perturbation lemma (see, e.g., [31, Lemma 2.3.2]) we deduce that $J_{\beta}(u)$ is invertible for every $u \in \mathcal{B}_{\delta_{\beta}}(\varphi(\beta)) \cap K_{\beta}$ and we obtain the following upper bound:

$$\left\| J_{\beta}(u)^{-1} J_{\varphi(\beta)} \right\|_{\infty} \leq \frac{4}{3}. \quad (24)$$

In order to describe our method, we need a sufficient condition for the convergence of the standard Newton iteration associated to (7) for any $\beta \in [\beta_*, \beta^*]$. Arguing as in [31, 10.4.2] we deduce the following remark, which in particular implies that the Newton iteration under consideration converges.

Remark 15. Set $\delta := \min\{\delta_\beta : \beta \in [\beta_*, \beta^*]\}$. Fix $\beta \in [\beta_*, \beta^*]$ and consider the Newton iteration

$$u^{(k+1)} = u^{(k)} - J_\beta(u^{(k)})^{-1}F(\beta, u^{(k)}) \quad (k \geq 0),$$

starting at $u^{(0)} \in K_\beta$. If $\|u^{(0)} - \varphi(\beta)\|_\infty < \delta$, then

$$\|u^{(k)} - \varphi(\beta)\|_\infty < \frac{\delta}{3^k}$$

holds for $k \geq 0$.

Now we can describe our homotopy continuation method. Let $\beta_0 := \beta_* < \beta_1 < \dots < \beta_N := \beta^*$ be a uniform partition of the interval $[\beta_*, \beta^*]$, with N to be fixed. We define an iteration as follows:

$$u^{(k+1)} = u^{(k)} - J_{\beta_k}(u^{(k)})^{-1}F(\beta_k, u^{(k)}) \quad (0 \leq k \leq N-1), \quad (25)$$

$$u^{(N+k+1)} = u^{(N+k)} - J_{\beta^*}(u^{(N+k)})^{-1}F(\beta^*, u^{(N+k)}) \quad (k \geq 0). \quad (26)$$

In order to see that the iteration (25)–(26) yields an approximation of the positive solution $\varphi(\beta^*)$ of (17) for $B = \beta^*$, it is necessary to obtain a condition assuring that (25) yields an attraction point for the Newton iteration (26). This relies on a suitable choice for N , which we now discuss.

By Theorem 14, we have

$$\begin{aligned} \|\varphi(\beta_{i+1}) - \varphi(\beta_i)\|_\infty &\leq \max\{\|\varphi'(\beta)\|_\infty : \beta \in [\beta_*, \beta^*]\} |\beta_{i+1} - \beta_i| \\ &\leq \kappa_1 \frac{\beta^*}{N}, \end{aligned}$$

for $0 \leq i \leq N-1$, where κ_1 is an upper bound of the condition number independent of h . Thus, for $N := \lceil 3\beta^*\kappa_1/\delta \rceil + 1 = O(1)$, by the previous estimate we obtain the following inequality:

$$\|\varphi(\beta_{i+1}) - \varphi(\beta_i)\|_\infty < \frac{\delta}{3} \quad (27)$$

for $0 \leq i \leq N-1$. Our next result shows that this implies the desired result.

Lemma 16. Set $N := \lceil 3\beta^*\kappa_1/\delta \rceil + 1$. Then, for every $u^{(0)}$ with $\|u^{(0)} - \varphi(\beta_*)\|_\infty < \delta$, the point $u^{(N)}$ defined in (25) is an attraction point for the Newton iteration (26).

Proof. By hypothesis, we have $\|u^{(0)} - \varphi(\beta_*)\|_\infty < \delta$. Arguing inductively, suppose that $\|u^{(k)} - \varphi(\beta_k)\|_\infty < \delta$ holds for a given $0 \leq k < N$. By Remark 15 we have that $u^{(k)}$ is an attraction point for the Newton iteration associated to (17) for $B = \beta_k$. Furthermore, Remark 15 also shows that $\|u^{(k+1)} - \varphi(\beta_k)\|_\infty < \delta/3$ holds. Then

$$\begin{aligned} \|u^{(k+1)} - \varphi(\beta_{k+1})\|_\infty &\leq \|u^{(k+1)} - \varphi(\beta_k)\|_\infty + \|\varphi(\beta_k) - \varphi(\beta_{k+1})\|_\infty \\ &< \frac{1}{3}\delta + \frac{1}{3}\delta < \delta, \end{aligned}$$

where the inequality $\|\varphi(\beta_{k+1}) - \varphi(\beta_k)\|_\infty < \delta/3$ follows by (27). This completes the inductive argument and shows in particular that $u^{(N)}$ is an attraction point for the Newton iteration (26). \square

Next we consider the convergence of (26), starting with a point $u^{(N)}$ satisfying the condition $\|u^{(N)} - \varphi(\beta^*)\|_\infty < \delta \leq \delta_{\beta^*}$. Combining this inequality with (22) we deduce that $u^{(N)} \in K_{\alpha^*}$. Furthermore, we see that

$$\begin{aligned} \|u^{(N+1)} - \varphi(\beta^*)\|_\infty &= \|u^{(N)} - J_{\beta^*}(u^{(N)})^{-1}F(\beta^*, u^{(N)}) - \varphi(\beta^*)\|_\infty \\ &= \left\| J_{\beta^*}(u^{(N)})^{-1} (J_{\beta^*}(u^{(N)})(u^{(N)} - \varphi(\beta^*)) - F(\beta^*, u^{(N)}) + F(\beta^*, \varphi(\beta^*))) \right\|_\infty \\ &\leq \|J_{\beta^*}(u^{(N)})^{-1} J_{\varphi(\beta^*)}\|_\infty \\ &\quad \left\| J_{\varphi(\beta^*)}^{-1} (J_{\beta^*}(u^{(N)})(u^{(N)} - \varphi(\beta^*)) - F(\beta^*, u^{(N)}) + F(\beta^*, \varphi(\beta^*))) \right\|_\infty \\ &\leq \|J_{\beta^*}(u^{(N)})^{-1} J_{\varphi(\beta^*)}\|_\infty \|J_{\varphi(\beta^*)}^{-1} (J_{\beta^*}(u^{(N)}) - J_{\beta^*}(\xi))\|_\infty \|u^{(N)} - \varphi(\beta^*)\|_\infty, \end{aligned} \tag{28}$$

where ξ is a point in the segment joining the points $u^{(N)}$ and $\varphi(\beta^*)$. Combining (21) and (24) we deduce that

$$\begin{aligned} \|u^{(N+1)} - \varphi(\beta^*)\|_\infty &< \frac{4}{3} \|J_{\varphi(\beta^*)}^{-1} (J_{\beta^*}(u^{(N)}) - J_{\beta^*}(\xi))\|_\infty \delta_{\beta^*} \\ &< \frac{4c}{3} \delta_{\beta^*}^2 \leq \frac{1}{3} \delta_{\beta^*} \end{aligned}$$

holds, with $c := (2\eta_{\beta^*}(\theta^* + 1)C(\beta^*)) / (g'_1(g^{-1}(C(\beta^*)/\beta^*))(1 - \rho^*))$. By an inductive argument we conclude that the iteration (26) is well-defined and converges to the positive solution $\varphi(\beta^*)$ of (17) for $B = \beta^*$. Furthermore, we conclude that the point $u^{(N+k)}$, obtained from the point $u^{(N)}$ above after k steps of the iteration (26), satisfies the estimate

$$\|u^{(N+k)} - \varphi(\beta^*)\|_\infty \leq \hat{c} \left(\frac{4c}{3} \delta_{\beta^*} \right)^{2^k} \leq \hat{c} \left(\frac{1}{3} \right)^{2^k},$$

with $\hat{c} := 3/4c$. Therefore, in order to obtain an ε -approximation of $\varphi(\beta^*)$, we have to perform $\log_2 \log_3(3/4c\varepsilon)$ steps of the iteration (26). Summarizing, we have the following result.

Lemma 17. *Let $\varepsilon > 0$ be given. Then, for every $u^{(N)} \in (\mathbb{R}_{>0})^n$ satisfying the condition $\|u^{(N)} - \varphi(\beta^*)\|_\infty < \delta$, the iteration (26) is well-defined and the estimate $\|u^{(N+k)} - \varphi(\beta^*)\|_\infty < \varepsilon$ holds for $k \geq \log_2 \log_3(3/4c\varepsilon)$.*

Let $\varepsilon > 0$. Assume that we are given $u^{(0)} \in (\mathbb{R}_{>0})^n$ such that $\|u^{(0)} - \varphi(\beta_*)\|_\infty < \delta$ holds. In order to compute an ε -approximation of the positive solution $\varphi(\beta^*)$ of (17) for $B = \beta^*$, we perform N iterations of (25) and $k_0 := \lceil \log_2 \log_3(3/4c\varepsilon) \rceil$ iterations of (26). From Lemmas 16 and 17 we conclude that the output $u^{(N+k_0)}$ of this procedure satisfies the condition $\|u^{(N+k_0)} - \varphi(\beta^*)\|_\infty < \varepsilon$. Observe that the Jacobian matrix $J_\beta(u)$ is tridiagonal for every $\beta \in [\beta_*, \beta^*]$ and every $u \in K_\beta$. Therefore, the solution of a linear system with matrix $J_\beta(u)$ can be obtained with $O(n)$ flops. This implies that each iteration of both (25) and (26) requires $O(n)$ flops. In conclusion, we have the following result.

Proposition 18. *Let $\varepsilon > 0$ and $u^{(0)} \in (\mathbb{R}_{>0})^n$ with $\|u^{(0)} - \varphi(\beta_*)\|_\infty < \delta$ be given, where δ is defined as in Remark 15. Then the output of the iteration (25)–(26) is an ε -approximation of the positive solution $\varphi(\beta^*)$ of (17) for $B = \beta^*$. This iteration can be computed with $O(Nn + k_0n) = O(n \log_2 \log_2(1/\varepsilon))$ flops.*

Finally, we exhibit a starting point $u^{(0)} \in (\mathbb{R}_{>0})^n$ satisfying the condition of Proposition 18. Let $\beta_* > 0$ be a constant independent of h to be determined. We study the constant

$$\delta := \min\{\delta_\beta : \beta \in [\beta_*, \beta^*]\},$$

where

$$\delta_\beta := \min \left\{ \frac{g'_1(g^{-1}(C(\beta)/\beta))(1 - \rho^*)}{8\eta_\beta(\theta^* + 1)C(\beta)}, g^{-1}(1/\beta) \right\}.$$

Since g_1 and g_2 are analytic functions in $x = 0$, in a neighborhood of $0 \in \mathbb{R}^n$, we can rewrite η_β as follows:

$$\begin{aligned} \eta_\beta &= 2 \max\{g''_1(2g^{-1}(1/\beta)), g''_2(2g^{-1}(1/\beta))g(g^{-1}(1/\beta))\} \\ &= 2 \max\{S_1(\beta), S_2(\beta)\}(g^{-1}(1/\beta))^{p-2}, \end{aligned}$$

where p is the multiplicity of 0 as a root of g_1 and S_i is an analytic function in $x = 0$ such that $\lim_{\beta \rightarrow 0} S_i(\beta) \neq 0$ for $i = 1, 2$. Taking into account that $\beta \in (0, \beta^*]$ holds, we conclude that there exists a constant $\eta^* > 0$, which depends only on β^* , with

$$\eta_\beta \leq 2\eta^*(g^{-1}(1/\beta))^{p-2}$$

for all $\beta \in (0, \beta^*]$. Moreover, with a similar argument we deduce that there exists a constant $\vartheta^* > 0$, which depends only on β^* , such that

$$\delta_\beta \geq \min \left\{ \frac{\vartheta^*(1 - \rho^*)}{16\eta^*(\theta^* + 1)C(\beta)} \left(\frac{g^{-1}(C(\beta)/\beta)}{g^{-1}(1/\beta)} \right)^{p-1}, 1 \right\} g^{-1}(1/\beta). \quad (29)$$

We claim that

$$\lim_{\beta \rightarrow 0^+} \frac{g^{-1}(C(\beta)/\beta)}{g^{-1}(1/\beta)} = 1^-. \quad (30)$$

In fact, since we have $C(\beta) \geq 1$ and g^{-1} is decreasing, it follows that

$$\frac{g^{-1}(C(\beta)/\beta)}{g^{-1}(1/\beta)} \leq 1. \quad (31)$$

On the other hand, there exists $\xi \in (1/\beta, C(\beta)/\beta)$ with

$$\begin{aligned} \frac{g^{-1}(C(\beta)/\beta)}{g^{-1}(1/\beta)} &= \frac{g^{-1}(1/\beta) + (g^{-1})'(\xi)((C(\beta) - 1)/\beta)}{g^{-1}(1/\beta)} \\ &= 1 + \frac{g(g^{-1}(1/\beta))}{g'(g^{-1}(\xi))g^{-1}(1/\beta)}(C(\beta) - 1) \\ &\geq 1 + \frac{g(g^{-1}(1/\beta))}{g'(g^{-1}(1/\beta))g^{-1}(1/\beta)}(C(\beta) - 1). \end{aligned} \quad (32)$$

Since g_1 and g_2 are analytic functions in $x = 0$ and $\lim_{\beta \rightarrow 0^+} C(\beta) = 1$, we see that

$$\lim_{\beta \rightarrow 0^+} \frac{g(g^{-1}(1/\beta))}{g'(g^{-1}(1/\beta))g^{-1}(1/\beta)}(C(\beta) - 1) = 0.$$

Combining (31), (32) and this inequality we immediately deduce (30).

Combining (29) with (30) we conclude that there exists a constant $C^* \in (0, 1]$, which depends only on β^* , with

$$\delta_\beta \geq C^* g^{-1}(1/\beta).$$

Therefore,

$$\delta = \min\{\delta_\beta : \beta \in [\beta_*, \beta^*]\} \geq C^* g^{-1}(1/\beta_*). \quad (33)$$

From Corollary 8 and Lemma 10, we have

$$\varphi(\beta_*) \in [g^{-1}(C(\beta_*)/\beta_*), g^{-1}(1/\beta_*)]^n.$$

Furthermore, by (30), we deduce that

$$\left(1 - \frac{g^{-1}(C(\beta_*)/\beta_*)}{g^{-1}(1/\beta_*)}\right) g^{-1}(1/\beta_*) < C^* g^{-1}(1/\beta_*) \quad (34)$$

holds for $\beta_* > 0$ small enough. Combining this with (33), we conclude that

$$\|u - \varphi(\beta_*)\|_\infty \leq g^{-1}(1/\beta_*) - g^{-1}(C(\beta_*)/\beta_*) < \delta \quad (35)$$

holds for all $u \in [g^{-1}(C(\beta_*)/\beta_*), g^{-1}(1/\beta_*)]^n$. Thus, let $\beta_* < \beta^*$ satisfy (34). Then, for any $u^{(0)}$ in the hypercube $[g^{-1}(C(\beta_*)/\beta_*), g^{-1}(1/\beta_*)]^n$, the inequality

$$\|u^{(0)} - \varphi(\beta_*)\|_\infty < \delta$$

holds. Therefore, applying Proposition 18, we obtain our main result.

Theorem 19. *Let $\varepsilon > 0$ be given. Then we can compute an ε -approximation of the positive solution of (17) for $B = \beta^*$ with $O(n \log_2 \log_2(1/\varepsilon))$ flops.*

References

- [1] J. Cannon, The one-dimensional heat equation, Cambridge Univ. Press, Cambridge, 1984.
- [2] C. Pao, Nonlinear parabolic and elliptic equations, Plenum Press, 1992.
- [3] J. Bebernes, D. Eberly, Mathematical problems from combustion theory, Vol. 83 of Applied Mathematical Sciences, Springer-Verlag, New York, 1989.
- [4] P. Grindrod, The theory and applications of reaction-diffusion equations: patterns and waves, Clarendon Press, Oxford, 1996.
- [5] J. Murray, Mathematical biology. Vol. 1: An introduction, Vol. 17 of Interdisciplinary Applied Mathematics, Springer, New York, 2002.
- [6] B. Gilding, R. Kersner, Travelling waves in nonlinear diffusion-convection reaction, Birkhäuser, Berlin, 2004.
- [7] H. Levine, The role of critical exponents in blow up theorems, SIAM Rev. 32 (1990) 262–288.
- [8] A. Samarskii, V. Galaktionov, S. Kurdyumov, A. Mikhailov, Blow-up in quasilinear parabolic equations, Vol. 19 of de Gruyter Expositions in Mathematics, Walter de Gruyter, Berlin, 1995.
- [9] M. Chipot, M. Fila, P. Quittner, Stationary solutions, blow up and convergence to stationary solutions for semilinear parabolic equations with nonlinear boundary conditions, Acta Math. Univ. Comenian. (N.S.) 60 (1) (1991) 35–103.

- [10] J. Lopez Gomez, V. Marquez, N. Wolanski, Dynamic behaviour of positive solutions to reaction–diffusion problems with nonlinear absorption through the boundary, *Rev. Un. Mat. Argentina* 38 (1993) 196–209.
- [11] P. Quittner, On global existence and stationary solutions for two classes of semilinear parabolic problems, *Comment. Math. Univ. Carolin.* 34 (1) (1993) 105–124.
- [12] J. Rossi, The blow-up rate for a semilinear parabolic equation with a nonlinear boundary condition, *Acta Math. Univ. Comenian. (N.S.)* 67 (2) (1998) 343–350.
- [13] J. Fernandez Bonder, J. Rossi, Blow-up vs. spurious steady solutions, *Proc. Amer. Math. Soc.* 129 (1) (2001) 139–144.
- [14] A. Rodríguez Bernal, A. Tajdine, Nonlinear balance for reaction–diffusion equations under nonlinear boundary conditions: Dissipativity and blow–up, *J. Differential Equations* 169 (2) (2001) 332–372.
- [15] F. Andreu, J. M. Mazón, J. Toledo, J. D. Rossi, Porous medium equation with absorption and a nonlinear boundary condition, *Nonlinear Anal.* 49 (2002) 541–563.
- [16] M. Chipot, P. Quittner, Equilibria, connecting orbits and a priori bounds for semilinear parabolic equations with nonlinear boundary conditions, *J. Dynam. Differential Equations* 16 (2004) 91–138.
- [17] C. Bandle, H. Brunner, Blowup in diffusion equations: a survey, *J. Comput. Appl. Math.* 97 (1998) 3–22.
- [18] E. Dratman, G. Matera, On the solution of the polynomial systems arising in the discretization of certain ODEs, *Computing* 85 (4) (2009) 301–337.
- [19] E. Dratman, Approximation of the solution of certain nonlinear ODEs with linear complexity, *J. Comput. Appl. Math.* 233 (9) (2010) 2339–2350.
- [20] M. De Leo, E. Dratman, G. Matera, Numeric vs. symbolic homotopy algorithms in polynomial system solving: A case study, *J. Complexity* 21 (4) (2005) 502–531.
- [21] L. Pardo, Universal elimination requires exponential running time, in: A. Montes (Ed.), *Computer Algebra and Applications, Proceedings of EACA–2000, Barcelona, Spain, September 2000*, 2000, pp. 25–51.

- [22] D. Castro, M. Giusti, J. Heintz, G. Matera, L. Pardo, The hardness of polynomial equation solving, *Found. Comput. Math.* 3 (4) (2003) 347–420.
- [23] E. Dratman, G. Matera, A. Weissbein, Robust algorithms for generalized Pham systems, *Comput. Complexity* 18 (1) (2009) 105–154.
- [24] E. Allgower, K. Georg, Numerical continuation methods: An Introduction, Vol. 13 of Series in Computational Mathematics, Springer Verlag, Heidelberg, 1990.
- [25] E. Allgower, D. Bates, A. Sommese, C. Wampler, Solution of polynomial systems derived from differential equations, *Computing* 76 (1–2) (2006) 1–10.
- [26] J. Duvallet, Computation of solutions of two-point boundary value problems by a simplicial homotopy algorithm, in: E. Allgower, K. Georg (Eds.), *Computational Solution of Nonlinear Systems of Equations*, Vol. 26 of Lectures in Appl. Math., Amer. Math. Soc., Providence, RI, 1990, pp. 135–150.
- [27] L. T. Watson, Solving finite difference approximations to nonlinear two-point boundary value problems by a homotopy method, *SIAM J. Sci. Stat. Comput.* 1 (4) (1980) 467–480.
- [28] B. Kacewicz, Complexity of nonlinear two-point boundary-value problems, *J. Complexity* 18 (3) (2002) 702–738.
- [29] L. Blum, F. Cucker, M. Shub, S. Smale, *Complexity and Real Computation*, Springer, New York Berlin Heidelberg, 1998.
- [30] G. Malajovich, J. Rojas, High probability analysis of the condition number of sparse polynomial systems, *Theoret. Comput. Sci.* 315 (2–3) (2004) 525–555.
- [31] J. Ortega, W. Rheinboldt, *Iterative solutions of nonlinear equations in several variables*, Academic Press, New York, 1970.